



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### Wave Directional Spreading From Point Field Measurements

**Citation for published version:**

McAlister, M, Venugopal, V & Borthwick, A 2017, 'Wave Directional Spreading From Point Field Measurements', *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 473, no. 2200, 20160781. <https://doi.org/10.1098/rspa.2016.0781>

**Digital Object Identifier (DOI):**

[10.1098/rspa.2016.0781](https://doi.org/10.1098/rspa.2016.0781)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



Research



**Cite this article:** McAllister ML, Venugopal V, Borthwick AGL. 2017 Wave directional spreading from point field measurements. *Proc. R. Soc. A* **473**: 20160781. <http://dx.doi.org/10.1098/rspa.2016.0781>

Received: 18 October 2016

Accepted: 21 March 2017

**Subject Areas:**

ocean engineering, oceanography, wave motion

**Keywords:**

surface gravity waves, directional spreading estimation, point measurement, second-order difference, *in situ* wave data

**Author for correspondence:**

M. L. McAllister

e-mail: [m.mcallister@ed.ac.uk](mailto:m.mcallister@ed.ac.uk)

# Wave directional spreading from point field measurements

M. L. McAllister, V. Venugopal and A. G. L. Borthwick

Institute for Energy Systems, University of Edinburgh, Kings Buildings, Edinburgh EH9 3JG, UK

MLM, 0000-0002-5142-3172

Ocean waves have multidirectional components. Most wave measurements are taken at a single point, and so fail to capture information about the relative directions of the wave components directly. Conventional means of directional estimation require a minimum of three concurrent time series of measurements at different spatial locations in order to derive information on local directional wave spreading. Here, the relationship between wave nonlinearity and directionality is utilized to estimate local spreading without the need for multiple concurrent measurements, following Adcock & Taylor (Adcock & Taylor 2009 *Proc. R. Soc. A* **465**, 3361–3381. (doi:10.1098/rspa.2009.0031)), with the assumption that directional spreading is frequency independent. The method is applied to measurements recorded at the North Alwyn platform in the northern North Sea, and the results compared against estimates of wave spreading by conventional measurement methods and hindcast data. Records containing freak waves were excluded. It is found that the method provides accurate estimates of wave spreading over a range of conditions experienced at North Alwyn, despite the noisy chaotic signals that characterize such ocean wave data. The results provide further confirmation that Adcock and Taylor's method is applicable to metocean data and has considerable future promise as a technique to recover estimates of wave spreading from single point wave measurement devices.

## 1. Introduction

The relative directions in which wave components travel have a substantial effect on their combined kinematics and dynamics. This has significant implications in coastal

and ocean engineering, including design codes for offshore oil platforms, ship safety, optimization of wave energy devices and the dispersion of oil slicks and plastic contaminants. For the majority of wave observations made in the oceans, the relative directions of the underlying wave components remain unknown because the observations are usually made in isolation as point measurements. The conventional way of overcoming this problem is to use either an array of three or more instruments [1,2] or multi-degree-of-freedom devices [3–5] such as radar, clover-leaf buoys and wave-rider buoys.

Another approach is to exploit the relationship between wave nonlinearity and wave directionality. The majority of waves in the ocean are weakly nonlinear; therefore, their observed time series will contain second-order nonlinearities [6,7]. The connection that exists between nonlinearity and directionality as observed experimentally by Johannessen & Swan [8], Onorato *et al.* [9], and in the field by Forristall [10], and Toffoli *et al.* [11], provides a means by which information on local directional spreading can be inferred from the nonlinearities that lie within a single point measurement. To exploit this connection, it is necessary to extract the linear free waves and second-order bound waves from the wave elevation time signal, and then fit second-order wave theory to the bound wave signal using an assumed spreading function, and hence estimate the spreading from the optimal fit. Walker *et al.* [12] presented an approximate method for separating out the linear free waves and second-order bound waves contained within an observed free surface elevation time series, through filtering in the frequency domain and minimizing skewness. For a prescribed wave spreading function involving a combination of interacting freely propagating linear waves, the resulting bound waves can be calculated to second-order using interaction kernels derived for finite depth by Dalzell [13]. Using this theory, the bound waves associated with a given set of free waves are simply a function of their component directions. By comparing the predicted bound second-order difference waves for an assumed spreading distribution to the actual bound second-order difference waves contained within the original measurement, Adcock & Taylor [14] derived a method for estimating the local directional spreading for a point measurement. This approach, herein called the ‘long-wave method’ (LWM), was satisfactorily validated using experimental data obtained by Cornett *et al.* [15] and numerically generated data with the addition of random noise. Adcock *et al.* [16] used the same approach to infer the conditions that gave rise to the Draupner wave of 1 January 1995 from information acquired by a single point gauge, and found that directional spreading was of key importance in understanding the extreme event. The results were found to be consistent with other nearby observations [17,18].

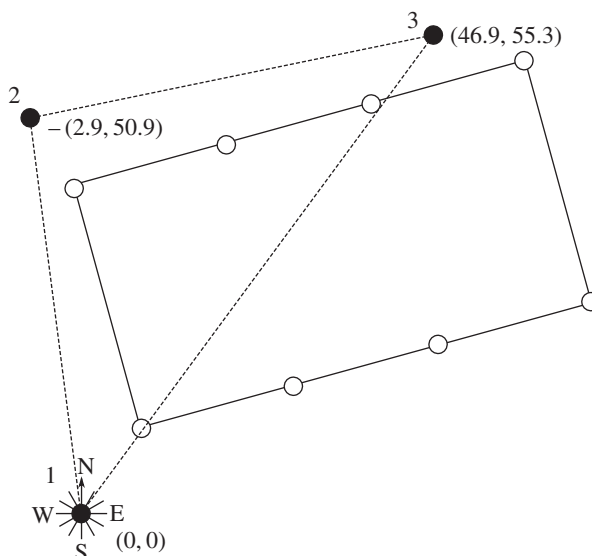
Here, Adcock and Taylor’s method is applied to a large dataset of observations from three measurement devices located on the North Alwyn platform in the northern North Sea. By examining high-quality data signals from all three devices, comparison is made between estimates of directional spreading using the single probe LWM and conventional methods. Where data are of insufficient quality from any of the three devices, comparison is drawn with hindcast data obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF).

The paper is structured as follows. Section 2 describes the data sources. Section 3 presents the methodology. Section 4 investigates the sensitivity of the method. Section 5 details the results. Section 6 lists the conclusions.

## 2. Data

### (a) North Alwyn platform

The ocean wave data considered herein were all obtained from single point gauges located on the North Alwyn platform, a fixed jacket offshore structure located in the northern North Sea at (60°48.5' N, 1°44.17' E) about 150 km east of the Shetland Isles. Mean water depth at the platform was 130 m. The platform comprised two structures, North Alwyn A and North Alwyn B, connected by a bridge. The platform substructure was sparse, with each support column of



**Figure 1.** Schematic showing arrangement of North Alwyn measurement array in plan.

diameter  $D = 1.5$  m. The wavelength of incoming waves was typically of the order of  $L = 100$  m. Hence,  $D/L \ll 0.2$  and the North Alwyn platform can be considered as being composed of small diameter cylinders, meaning that wave–structure interactions such as reflections and diffraction were negligible [19] and did not significantly affect the recorded observations of the surrounding wave field.

The measurement devices were all located on North Alwyn A, and comprised three Thorn EMI infrared laser probes that simultaneously measured sea surface elevations (figure 1). The probes were set out in plan as nodes of a triangle with side lengths of about 50, 51 and 72.5 m. The probe resolution was  $\pm 5$  cm (i.e. accurate to within 1%). It should be noted that the water elevation time series measured by the probes contained a number of wild points and occasional signal drop-out exhibited as plateaux in the data [20]. A Labtech Notebook on the platform was used to control data acquisition, undertake preliminary data processing, and provide local data storage. The PC acquired data at 5 Hz, via an XE software package. Raw data were split into 20 min blocks for statistical treatment. Further details are given by Wolfram *et al.* [21].

Free surface elevation data were recorded over a 10 year period from 1994 to 2004. The data measurements were continuously monitored, and values stored when either the significant wave height  $H_s$  exceeded 3.5 m or the wind speed was greater than  $16 \text{ ms}^{-1}$ . A total of 448 individual ‘storms’ made up the North Alwyn dataset, comprising 16 422 separate 20 min duration files, each containing three concurrent sets of measurements. Prior to analysis, data were screened using a strict quality control process based on a procedure outlined by Christou & Ewans [20,22]. All data that passed these tests were subjected to further analysis. Directional analysis was performed by conventional methods when the measurements from all three probes were simultaneously of good quality. Table 1 summarizes these results, listing the number of measurements that passed the quality control process. Of the unprocessed measurements, 20 802 from any of the three sensors passed the quality control procedure, and of these 391 were suitable for multi-probe directional analysis.

## (b) Hincast data from European Centre for Medium Range Weather Forecasts

The ECMWF provides publicly available meteorological datasets produced by hindcast models. The data used herein were taken from the ERA-20C model, which provided hindcast atmospheric

**Table 1.** North Alwyn Dataset: availability of measurements between 1994 and 2005, and results of quality control process.

year	storms	files	QC passed	3 probes QC passed
1994	1	177	172	124
1995	5	658	595	0
1996	1	71	148	22
1997	5	1241	1659	0
2000	176	5972	7698	185
2001	100	2272	2581	12
2002	108	4019	4792	44
2003	52	2012	3157	4
total	448	16 422	20 802/49 266	391/16 422

information based on an assimilation of historic global observations for the entire twentieth century. Model outputs included a wide range of meteorological data; of these, the key metocean parameters were hindcast using a WAM model. The WAM model outputs spectra, discretized at 36 directions and frequencies. The ‘wave spectral directional width’  $\sigma_E$  output had a spatial resolution of  $\delta x = 28$  km and temporal resolution of  $\delta t = 3$  h. A full specification of the wave model is given by ECMWF [23].

### 3. Method

#### (a) Linearization

In order to calculate the bound waves for a given time series of free surface elevation above the mean water level  $\eta(t)$ , the linear free waves  $\eta_L$  must first be extracted. The data were linearized following the procedure outlined by Walker *et al.* [12]. Second-order difference waves  $\eta_{2F-}$  were removed by high-pass filtering the data at half the peak frequency. Second-order sum nonlinearity was then removed by adjusting the Stokes’  $S_{22}$  parameter to minimize the skewness of the resulting linear time series.

$$\eta_L = \eta_F - \frac{S_{22}}{h}(\eta^2 - \hat{\eta}^2), \quad (3.1)$$

where  $h$  is the mean water depth, and  $\eta_L$  is the linearized,  $\eta_F$  is the high-pass filtered, and  $\hat{\eta}$  is the Hilbert-transformed, time series obtained from  $\eta$ . The skewness of  $\eta_L$  is calculated using

$$\gamma = \frac{1}{N\sigma^3} \sum_{n=1}^N \eta_L^3, \quad (3.2)$$

where  $N$  is the number of data points and  $\sigma$  is the standard deviation of the linearized time series.

#### (b) Second-order calculations

Wave-wave interaction of linear freely propagating waves causes nonlinear bound waves which occur at increasing order with decreasing amplitude. Second-order bound waves occur as a result of interacting wave pairs, at the sum and difference of their constituents. The sum terms occur in the tail of the spectrum and overlap the linear waves. However, the difference terms occur at the low end of the spectrum where they dominate linear waves, owing to the faster decay of the linear spectrum. This makes it easier to separate out the difference waves.

Difference waves resulting from a linear spectrum may be calculated using the interaction kernels derived by Dalzell [13] for finite depth. Starting with the linear spectrum expressed as a complex Fourier series

$$\eta = \sum_{n=1}^{n=N} x_n \exp(i\omega_n t), \quad (3.3)$$

in which  $x_n$  is the  $n$ th constituent of the complex vector  $\mathbf{x}$  generated by performing a fast Fourier transform (FFT) on the free surface time series,  $\omega_n$  is the corresponding angular frequency of the  $n$ th constituent, and  $t$  is time. Calculating the interaction of each wave pair gives the second-order difference waves

$$\eta_{2-} = \sum_{i=1}^{i=N} \sum_{j=1}^{j=N} \sum_{n=1}^{n=N} \sum_{m=1}^{m=N} d(\theta_i) x_n d(\theta_j) x_m K^- \exp(i(\omega_n - \omega_m)t), \quad (3.4)$$

where  $d$  is the assumed spreading function, dependent on wave incidence angle  $\theta$ , and  $K^-$  is the interaction kernel for difference terms.  $K^-$  is a function of  $\omega$ ,  $\theta$ ,  $h$ , and wavenumber  $k$ , and is defined as

$$\begin{aligned} K^- = & \frac{\omega_n^2 + \omega_m^2}{2g} + \frac{\omega_n \omega_m}{2g} \left( 1 + \frac{\cos(\theta_n - \theta_m)}{\tanh(|k_n|/h) \tanh(|k_m|/h)} \right) \\ & \times \left( \frac{(\omega_n - \omega_m)^2 + g|k_n - k_m| \tanh(|k_n - k_m|/h)}{D_m(k_n, k_m)} \right) \\ & + \frac{(\omega_n - \omega_m)}{2g D_m(k_n, k_m)} \left[ \frac{\omega_n^3}{\sinh^2(|k_n|/h)} - \frac{\omega_m^3}{\sinh^2(|k_m|/h)} \right] \end{aligned} \quad (3.5)$$

where

$$D_m(k_n, k_m) = (\omega_n - \omega_m)^2 - g|k_n - k_m| \tanh(|k_n - k_m|/h). \quad (3.6)$$

### (i) Numerical implementation

Equation (3.4) involves a quadruple summation over both frequency and direction. This incurs large computational cost, which is reduced using the following procedure. Dropping the exponential term in equation (3.4), the complex vector that represents the difference waves may be expressed as

$$\mathbf{x}_{2-} = \sum_{i=1}^{i=N} \sum_{j=1}^{j=N} \sum_{n=1}^{n=N} \sum_{m=1}^{m=N} d(\theta_i) x_n d(\theta_j) x_m K^-. \quad (3.7)$$

Noting that  $d(\theta)$  is the only independent variable in equation (3.7) for a given time series, then  $K^-$  can be expressed as an element of the  $n \times m$  matrix  $\mathbf{K}^-(\theta_\delta)$ , where  $\theta_\delta$  is simply the angle between interacting wave pairs. Hence, for frequency-independent cases  $x_n x_m$  and  $d(\theta_i) d(\theta_j)$  can be expressed in matrix form as

$$\mathbf{X}(n, m) = \mathbf{x}^T \mathbf{x}, \quad (3.8)$$

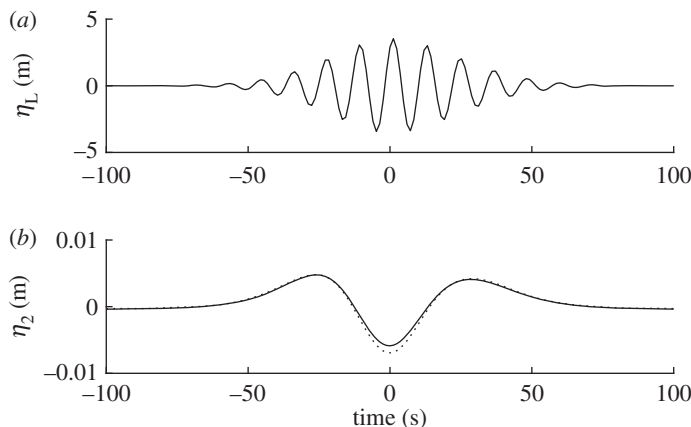
and

$$\mathbf{D}(i, j) = \mathbf{d}^T \mathbf{d}. \quad (3.9)$$

$\mathbf{D}$  can be transformed to a function of  $\theta_\delta = \theta_i - \theta_j$ , by summation over its diagonals

$$\mathbf{d}(\theta_\delta) = \sum_{i-j=0}^{i-j=N} D_{ij}. \quad (3.10)$$

Summing over all angles of separation gives the complex matrix  $\mathbf{X}_{2-}$ , which represents the interaction of all frequency pairs over all directional pairs. Diagonal elements of this matrix



**Figure 2.** NewWave verification: (a) linear focused wave group free surface elevation time series  $\eta$  at  $x = 0$ ; and (b) second-order difference wave time series showing interaction kernels  $\eta_{2I-}$  (solid line) and fully nonlinear potential flow  $\eta_{2F-}$  (dashed line) [24].

represent values of equal resultant frequency, therefore summation over diagonals provides the complex vector  $\mathbf{x}_{2-}$  from

$$\mathbf{x}_{2-} = \sum_{n-m=0}^{n-m=N} X_{2-nm}, \quad (3.11)$$

where

$$\mathbf{x}_{2-}(n, m) = \sum_{\delta=0}^{\delta=N} d(\theta_{\delta}) \mathbf{X} \mathbf{K}^{-}(\theta_{\delta}). \quad (3.12)$$

Finally, performing the inverse FFT on  $\mathbf{x}_{2-}$  gives the second-order difference waves

$$\eta_{2-} = \sum_{n=1}^{n=N} x_{2-n} \exp(i\omega_{2-n} t), \quad (3.13)$$

in which  $\omega_{2-n}$  is the  $n$ th difference frequency.

This method was tested against fully nonlinear potential flow simulations of a directionally spread focused NewWave group on water of infinite depth carried out by Gibbs & Taylor [24], using a numerical solver developed by Bateman *et al.* [25]. Figure 2 shows the linear free surface of the simulations and the corresponding bound difference waves obtained using the present method and that of Gibbs & Taylor [24]. Excellent agreement is obtained regarding the long waves; there is nevertheless a slight discrepancy at focus which is the result of narrowing of the directional spectrum due to nonlinearity.

## (c) Spreading estimation

### (i) Point measurement: long-wave method

Theoretical difference waves can be produced for any arbitrary spreading distribution using equation (3.5) to (3.13) for the linearized time series  $\eta_L$ . Here, the predicted difference waves,  $\eta_{2T-}$ , are compared with observed difference waves,  $\eta_{2F-}$  from (3.13), that were extracted by filtering the measured free surface time series  $\eta$ . By taking the Euclidean norm,

$$\Delta_{\text{LW}} = \frac{\sqrt{\sum (\eta_{2T-} - \eta_{2F-})^2}}{\sqrt{\sum (\eta_{2F-})^2}}. \quad (3.14)$$



The value of  $\Delta_{\text{LW}}$  is then minimized by varying the assumed spreading distribution. In this case, a wrapped-normal distribution is used, where the spreading value  $\sigma$  corresponds to standard deviation of a normal distribution wrapped around a full circle

$$d(\theta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\theta^2}{2\sigma^2}\right). \quad (3.15)$$

It is assumed that spreading is frequency independent. The value  $\sigma_{\text{opt}}$  that minimizes  $\Delta_{\text{LW}}$  and hence provides the optimum fit between the observed and theoretical difference waves  $\eta_{2\text{T-}}$  and  $\eta_{2\text{F-}}$  is determined as the best estimate of the local spreading.

By performing the foregoing LWM analysis on the entire dataset, any data for which the method did not successfully minimize the long-wave difference between  $0\text{--}90^\circ$  were discarded. Data where freak waves were located were also removed, because such waves have been shown to cause anomalous second-order difference waves [16]. Here, a freak wave is defined as a wave with up or down-crossing height greater than twice the significant wave height  $H_{m0}$  [26].

## (ii) Array measurements

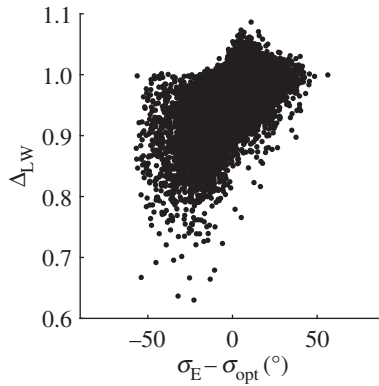
Where data of sufficiently high quality were simultaneously available from the three gauges at North Alwyn, the iterative maximum likelihood method (IMLM) is used to estimate the directional spectrum  $d(\theta)s(\omega)$ . From this spectrum, the spectrally weighted frequency-independent spreading distribution  $d(\theta)$  is calculated, and a wrapped-normal spreading distribution with standard deviation  $\sigma_{\text{D}}$  fitted to  $d(\theta)$ .

# 4. Sensitivity

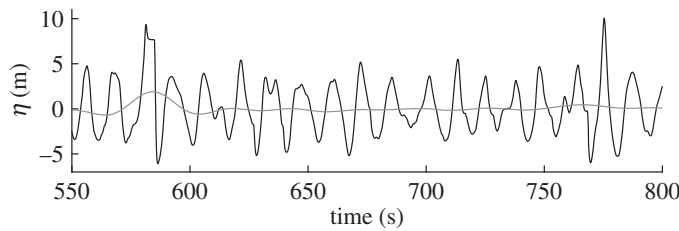
## (a) Goodness of fit

The long-wave difference  $\Delta_{\text{LW}}$  is primarily used as a goodness-of-fit parameter, in finding the optimal assumed spreading value  $\sigma_{\text{opt}}$ . However,  $\Delta_{\text{LW}}$  also provides information on the quality of fit that was achieved in finding  $\sigma_{\text{opt}}$ . High values of  $\Delta_{\text{LW}}$  corresponding to  $\sigma_{\text{opt}}$  imply that the fit may not be particularly satisfactory despite  $\sigma_{\text{opt}}$  being the best fit. Where a good fit is not achieved, the accuracy of  $\sigma_{\text{opt}}$  may be affected adversely. Figure 3 shows how  $\Delta_{\text{LW}}$  varies against the spreading error  $\sigma_{\text{E}} - \sigma_{\text{opt}}$ , where  $\sigma_{\text{E}}$  is the value of spreading predicted by the ECMWF hindcast. In this instance, hindcast data are used for mainly illustrative purposes; the validity of this comparison is discussed in §5a. However suitable the comparison may be, large differences in spreading are clearly indicative of error. For  $\Delta_{\text{LW}} \gtrsim 1$ , the error  $\sigma_{\text{opt}} - \sigma_{\text{E}} \rightarrow +$ , meaning  $\sigma_{\text{opt}} < \sigma_{\text{E}}$  and the spreading is underestimated. As  $\Delta_{\text{LW}} \rightarrow 0$ , the error also reduces. However, for  $\Delta_{\text{LW}} < 0.8$ , the spreading is overestimated. In such cases a seemingly good fit is achieved (as indicated by the very low value of  $\Delta_{\text{LW}}$ ) despite a large discrepancy in the predicted value of spreading. Visual inspection of the results indicated that this arose from erroneous measurements. Figure 4 shows an example of this type of error, where near flat sections or sudden changes in slope that are missed by the quality control process cause pronounced set-ups in the filtered second order difference waves  $\eta_{2\text{F-}}$ . In this example, there are two similarly sized waves in  $\eta(t)$  occurring at around 580 and 775 s, of crest height  $\sim 10$  m. Both waves cause set-ups in  $\eta_{2\text{F-}}$  as may be expected for large amplitude crests. The set-ups associated with the waves are of amplitude  $\sim 2$  m and  $\sim 0.5$  m, respectively. The flat section measured in the first wave is clearly an error, and not an artefact of the wave profile. This type of error typically occurs when the sensor loses the free surface temporarily and logs the previously recorded value until the surface is found once more. To assess the effect this error has on  $\eta_{2\text{F-}}$ , the flat error was artificially smoothed out as shown by the dashed line in figure 5a. Referring to the array shown in figure 1, the solid line corresponds to measurements at Probe 1 and the grey line to Probe 3 for comparison. Measurements at Probe 2 are omitted in figure 5 owing to poor quality.

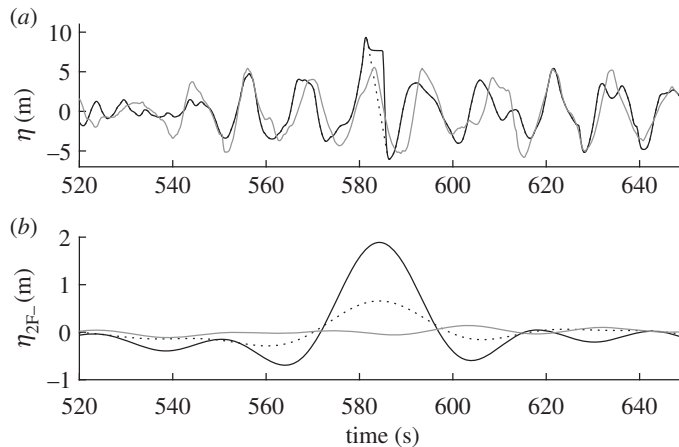




**Figure 3.** Scatter plot of goodness-of-fit estimates against spreading error for the LWM compared with the ECMWF hindcast.



**Figure 4.** Free surface elevation and second-order difference time series at North Alwyn for a selected storm where there is a large set-up in second-order waves owing to measurement error:  $\eta(t)$  free surface elevation (black); and  $\eta_{2F-}(t)$  filtered difference waves (grey).



**Figure 5.** Wave free surface elevation time series at North Alwyn showing unmodified signal (Probe 1, solid black line), smoothed signal (Probe 1, black dashed line) and concurrent measurement from Probe 3 (grey line): (a) free surface elevation time series,  $\eta(t)$ ; and (b) filtered second-order difference waves  $\eta_{2F-}(t)$ .

Linear interpolation was used to remove instrumentation error at the crest of the first large wave, and so the resulting shape of the wave measured at Probe 1 is slightly larger but not dissimilar in shape to the corresponding wave recorded at Probe 3 [figure 5a](#). Filtering the modified time series to obtain  $\eta_{2F-}$  results in the dashed line shown in [figure 5b](#); the black solid line shows

**Table 2.** Spreading estimates produced using raw and modified time series which correspond to three outlying estimates of  $\sigma_{\text{opt}}$  in Storm 369, that were a result of measurement errors.

outlier	raw		modified	
	$\sigma_{\text{opt}}(^{\circ})$	$\Delta_{\text{LW}}$	$\sigma_{\text{opt}}(^{\circ})$	$\Delta_{\text{LW}}$
(a)	76.6	0.942	39.6	0.971
(b)	46.4	0.915	31.1	0.989
(c)	52.4	0.839	28.1	0.996

$\eta_{2\text{F-}}$  for the unmodified time series. The modified time series has an amplitude of  $\sim 0.65$  m, which is comparable to that associated with the second large wave at 775 s.

The LWM spreading estimates from the modified and unmodified time series are  $34.6^{\circ}$  and  $86.3^{\circ}$ , respectively. The modified value concurs with the prediction made by the ECMWF hindcast of  $30.8^{\circ}$ , as well as the other values of  $\sigma_{\text{opt}}$  estimated during the rest of the storm.

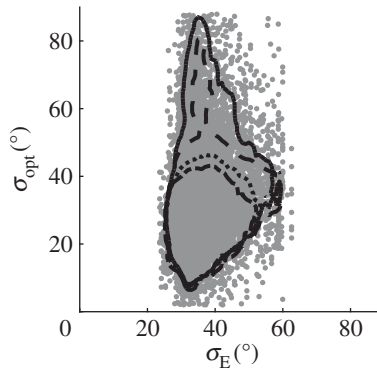
This approach was extended to Storm 369 of the North Alwyn dataset, discussed in more detail in §5b. This storm was selected because it possesses three outlying results that derive from obvious measurement errors that have passed the QC process. Additionally, 15 measurements taken during the storm allow for IMLM estimation of  $\sigma_{\text{D}}$ , which provides further confidence in the expected values of spreading. Again, available concurrent measurements were used as a guide when smoothing out measurement errors. Table 2 lists the results obtained after modifying the erroneous time series. The modified estimates of  $\sigma_{\text{opt}}$  lie much closer to the values of both  $\sigma_{\text{E}}$  and  $\sigma_{\text{D}}$  which indicate that the  $\sigma \approx 31^{\circ}$  for Storm 369. For all three cases, this relatively simple approach reduces the associated error in predicted spreading, despite reducing the goodness of fit. The interpolation method used to smooth out errors has a significant effect on the estimate of  $\sigma_{\text{opt}}$ . As there is no way of establishing for certain the true waveform, it is not suggested that this approach is used as a means of producing spreading estimates. However, the results presented here illustrate a particular source of sensitivity which does not lie with the LWM itself. It is obvious that errors undetected by the quality control process can result in gross overestimation of the spreading. Herein, visual inspection was used. In practice, a more robust means of error detection is desirable.

### (i) NewWave comparison

NewWave [27], depicted in figure 2a, constitutes a focused wave group where all free wave components are in phase at  $x = 0$  and  $t = 0$ . In practice, this waveform is used as a design wave because it provides an accurate approximation to the shape of the largest waves contained within ocean observations (e.g. [28]). Similarly, the second-order difference waves  $\eta_{2\text{NW-}}$  corresponding to this profile can be used to provide an approximation to those expected in large wave events [29]. The phasing and shape of actual waves affect the amplitude of  $\eta_{2\text{F-}}$ , causing some scatter about the NewWave approximation. Even so, this approach provides a general guide to the amplitude that should be expected for a given set of conditions. For a given spectral density  $S(\omega)$ , the corresponding new wave profile is calculated as

$$\eta_{\text{NW}}(t) = \frac{\sum_{n=1}^N S(\omega_n) \cos(\omega_n t)}{\sum_{n=1}^N S(\omega_n)}, \quad (4.1)$$

in which  $t$  is time from focus, located at  $x = 0$ . Using the linear NewWave profile  $\eta_{\text{NW}}$ , the second-order difference bound waves  $\eta_{2\text{NW-}}$  are calculated using equation (3.4). The second-order difference amplitude associated with the NewWave profile normalized by significant wave height  $a_{2\text{NW-}}/H_s$  can then be used to establish whether or not the observed set-up in amplitude  $a_{2\text{F-}}/H_s$  found in  $\eta_{2\text{F-}}$  is feasible. When normalized in this way, the value of  $a_{2\text{NW-}}/H_s$  depends upon the peak period  $T_p$  and spreading angle, with the assumption that the corresponding linear



**Figure 6.** Scatter diagram showing LWM estimate spreading values filtered using NewWave amplitude ratio,  $a_{2-}/a_{2NW-}$  plotted against hindcast estimates of spreading (grey dots), with superimposed contours showing 95% confidence values; results screened using amplitude ratio  $a_{2-}/a_{2NW-}$  values less than 100 (solid line), 25 (dashed line), less than 10 (dotted line) and less than 5 (dot-dashed line).

amplitude of the NewWave profile is  $a/H_s = 1$  for a JONSWAP spectrum with peak enhancement factor 3.3. Therefore, a NewWave amplitude can be calculated for each measurement using the appropriate values of  $T_p$  and  $\sigma_E$ . The ratio of  $a_{2F-}/H_s$  to  $a_{2NW-}/H_s$  provides an understanding of whether the set-up in  $\eta_{2F-}$  is likely to be the result of a large wave or arising from an anomaly. Figure 6 shows the effect of using  $a_{2F-}/a_{2NW-}$  as a quality control parameter. Values of  $\sigma_{opt}$  are plotted against the corresponding values of  $\sigma_E$ . Overlaid are the 95% contours when the data are screened using decreasing ratios of  $a_{2F-}/a_{2NW-}$  from 100 to 5. At high ratio of  $a_{2F-}/a_{2NW-}$  the amplitude observed in the measurement is much larger than would reasonably be expected; conversely as the ratio is reduced, results where the long wave estimate is grossly overestimated are progressively filtered out. This further illustrates the method's sensitivity to measurement errors. The results from the North Alwyn data confirm that NewWave provides an effective *a posteriori* method for error detection.

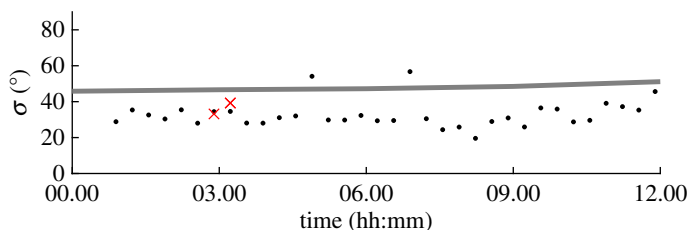
## (b) Cut-off frequency

Abnormal set-ups in the filtered difference waves,  $\eta_{2F-}$ , have been shown above to result from errors in the measured time series  $\eta$ . Moreover, if the linear spectrum does not decay as sharply as expected, large-amplitude linear components arise in  $\eta_{2F-}$  that can dominate the smaller-amplitude second-order components. The resulting increases in  $\eta_{2F-}$  amplitude may adversely affect the values estimated by the LWM. To overcome this, we reduce the frequency at which the original data  $\eta$  are filtered in order to obtain  $\eta_{2F-}$  which is uncontaminated by the effect of any linear components that lie close to the original frequency cut-off. This was examined by applying the LWM to the entire dataset with difference waves filtered at  $0.5f_p$ ,  $0.4f_p$  and  $0.3f_p$ . No difference in the estimated values of  $\sigma_{opt}$  was discernible when different frequency cut-offs were used.

## 5. Results

### (a) Correlation with European Centre for Medium-Range Weather Forecasts hindcast

For the majority of the data, measurement quality does not permit conventional directional estimation, and so an indirect means of comparison is necessary. Here, the 'wave spectral directional width'  $\sigma_E$  produced by the ECMWF ERA-20C model is used. The indirect nature and low resolution relative to that of the North Alwyn dataset means that direct comparison between the values of  $\sigma_{opt}$  and  $\sigma_E$  is of limited use on a measurement by measurement basis. However, as



**Figure 7.** Storm 50 of the North Alwyn dataset which shows a bias between ECMWF and LWM estimates of directional spreading:  $\sigma_{\text{opt}}$  (black dots);  $\sigma_E$  (grey line); and  $\sigma_D$  ( $\times$ ). (Online version in colour.)

opposed to comparing individual values, examination as to how the two sets of data correlate in time is much more informative. From figure 6, it is apparent that the values of spreading produced by the hindcast model are limited to the range of approximately  $20^\circ \leq \sigma_E \leq 60^\circ$ , this may be a result of the  $10^\circ$  directional resolution of the hindcast model. It is therefore not possible to assess values of  $\sigma_{\text{opt}} < 20^\circ$ , through this means of comparison.

The majority of storms in the North Alwyn dataset are relatively short in duration with an average length of 12 h, and consequently their spreading shows little temporal variation. For such storms, it is difficult to find correlation, and accordingly it is difficult to establish whether either  $\sigma_{\text{opt}}$  or  $\sigma_E$  correctly predict spreading if their values are different. Figure 7 shows an example of a storm where this is the case. Both the values of  $\sigma_{\text{opt}}$  and  $\sigma_E$  show little variation over the duration of the storm, with a gradual increase in spreading as the storm progresses and a mean difference of  $\sim 15^\circ$ , and it is difficult to draw any immediate conclusions. However, for this particular storm two of the measurement files were suitable for conventional estimation, as shown by the crosses. At about 03.00, the hindcast prediction  $\sigma_E = 46.6^\circ$  and the mean of two values either side of 03.00 are  $36.2^\circ$  and  $34.5^\circ$  for  $\sigma_D$  and  $\sigma_{\text{opt}}$ , respectively. This limited additional information suggests that the values of  $\sigma_{\text{opt}}$  are more likely to be correct, and the hindcast model overestimates the spreading.

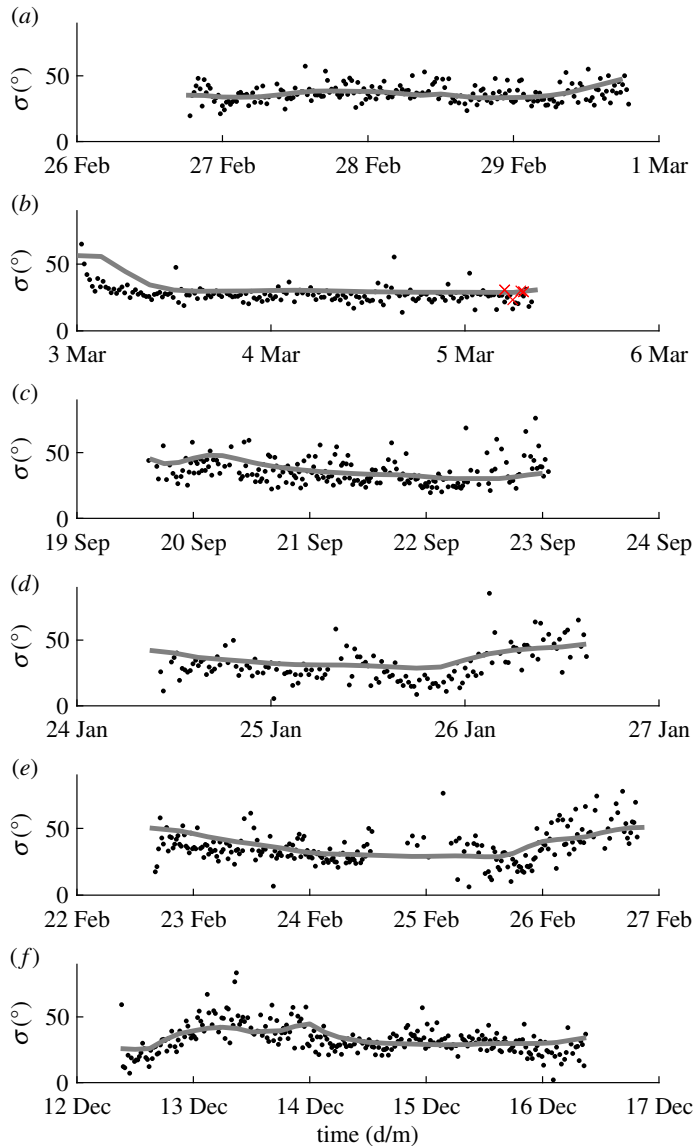
When there is variation in spreading predicted by the hindcast model it is possible to establish whether  $\sigma_{\text{opt}}$  is correctly measuring the local spreading. If a strong correlation exists between the values of  $\sigma_{\text{opt}}$  and  $\sigma_E$ , it is clear that the hindcast is modelling the same conditions as the LWM is measuring. A means of parametrizing the correlation observed between the two variables is the correlation coefficient. The covariance of two variables  $A$  and  $B$  is given by

$$\text{cov}(A, B) = \frac{1}{N-1} \sum_{i=1}^N (A_i - \mu_A)(B_i - \mu_B), \quad (5.1)$$

where  $\mu$  is the mean value of each signal, and  $N$  is the number of samples. The correlation coefficient  $\rho$  is then calculated by normalization, using the standard deviation of both signals,

$$\rho = \frac{\text{cov}(A, B)}{\sigma_A \sigma_B}. \quad (5.2)$$

Figure 8 shows the temporal behaviour of the ‘long wave estimate’ of spreading obtained for six storms chosen from the dataset. These storms were selected because they exhibited the strongest correlation between  $\sigma_{\text{opt}}$  and  $\sigma_E$  corresponding to the largest values of  $\rho$ , while having significant temporal variation over their duration. Effects of random uncertainty and/or noise are evident in the values of  $\sigma_{\text{opt}}$ . This variability stems from noise that is naturally found in real measurements, causing the standard deviation of the resulting estimates to increase. Adcock & Taylor [14] used numerically generated examples with artificially added noise to demonstrate that the standard deviation of the resulting values of  $\sigma_{\text{opt}}$  increased as the signal-to-noise ratio decreased. However, the mean value of the estimates remained correct.



**Figure 8.** LWM spreading estimate and ECMWF hindcast predictions with time. (a) Storm 28, (b) Storm 30, (c) Storm 134, (d) Storm 301, (e) Storm 320 and (f) Storm 435 as detailed in [table 3](#):  $\sigma_{\text{opt}}$  (black dots);  $\sigma_E$  (grey line); and where available  $\sigma_D$  ( $\times$ ). (Online version in colour.)

Noise contained within the values of  $\sigma_{\text{opt}}$  has an effect on the calculation of  $\rho$ . For two perfectly correlated signals where  $\rho$  is 1, the introduction of random errors invariably reduces the value of  $\rho$  [30]. To achieve a better understanding of the true correlation of the results, raw values of  $\sigma_{\text{opt}}$  were smoothed in the time domain with local regression using weighted linear least squares and a second degree polynomial model. [Table 3](#) lists values of  $\rho$  for both the smoothed and raw LWM estimates. For the raw results, the values of  $\rho$  are relatively low due to noise in the data. Even so,  $\rho \gtrsim 0.5$ , except for storms 28 and 134, meaning a weak correlation exists. After data smoothing, all the values of  $\rho$  are well above 0.5, with some approaching unity, illustrating a strong correlation between  $\sigma_{\text{opt}}$  and  $\sigma_E$ .

Not all of the storms in the dataset show such strong correlation. The indirect nature of the hindcast data means that negative comparisons are fairly inconclusive. However, when

**Table 3.** Fit parameters of well-correlated storms, correlation coefficient  $\rho$  for raw and smoothed estimates, and number of high-quality time series  $N_{QC}$ .

storm	start	end	$\rho$		$N_{QC}$
			raw	smoothed	
28	9 Feb 2000 08.58.28	14 Feb 2000 02.39.12	0.268	0.653	207
30	3 Mar 2000 02.14.06	5 Mar 2000 02.34.26	0.528	0.917	166
134	19 Sep 2000 14.47.46	22 Sep 2000 04.32.10	0.316	0.775	196
301	24 Jan 2002 10.21.18	26 Jan 2002 03.21.28	0.648	0.889	140
320	22 Feb 2002 16.09.18	26 Feb 2002 03.29.50	0.494	0.809	222
435	12 Dec 2003 09.12.14	16 Dec 2003 03.32.46	0.566	0.759	265

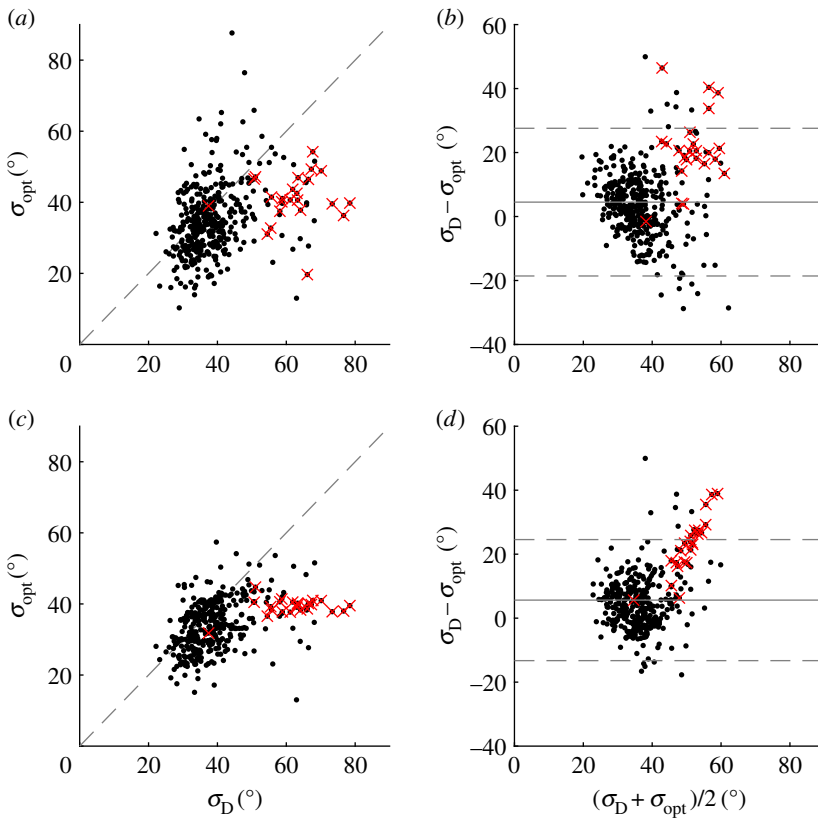
the hindcast and measurements exhibit high correlation, this provides conclusive evidence that the LWM is capturing the dynamically changing directional conditions being modelled by the hindcast. It is very unlikely that these two predictions correlate by pure coincidence without being an accurate measure of actual observed conditions.

### (b) Comparison with iterative maximum-likelihood method estimates

Of the 448 storms considered herein, 66 storms contain a total of 391 individual measurements for which all three probes pass the quality control process, allowing estimation of  $\sigma_D$ . This represents a very small portion of the entire dataset of 16 422 measurements. Unlike the hindcast data estimate  $\sigma_E$ , calculation of  $\sigma_D$  provides a direct measurement of local spreading. Figure 9a shows a direct comparison between the values of  $\sigma_D$  and  $\sigma_{opt}$  calculated for each set of concurrent measurements. The data are scattered about the line  $\sigma_D = \sigma_{opt}$ ; as the apparent value of spreading increases, the agreement between the two estimates appears to reduce. Figure 9b presents the Bland–Altman plot which displays the difference between  $\sigma_D$  and  $\sigma_{opt}$  against their average value. The Bland–Altman plot illustrates the uncertainty inherent to both sets of estimates and the relationship between magnitude and level of agreement, without assuming that one method is better than the other [31]. The mean difference for all data, shown by the solid grey line, is  $4.98^\circ$ ; this represents the bias between the two methods. As the mean value increases, the agreement between the two methods reduces, and the uncertainty increases. The dashed lines show 95% confidence intervals, which are calculated as twice the standard deviation from the mean; the majority of the data sit within these limits.

Given that the LWM requires only one measurement for each estimate, the number of available estimates means that it is possible to reduce the effects of random uncertainty by smoothing the data in the time domain. Smoothing is achieved using the same approach described in §5a. Figure 9c,d shows the effect of smoothing the LWM results on the agreement between the two methods. This removes some of the uncertainty associated with  $\sigma_{opt}$  (as addressed in the previous section) and has two effects. Firstly, the agreement increases for the majority of the data, as evident in figure 9d where the data are more tightly clustered about the mean. Secondly, it serves to highlight the uncertainty associated with the values of  $\sigma_D$ . As the mean value of spreading increases, agreement decreases with a positive bias, owing to the larger values of  $\sigma_D$ .

Considering the relatively small sample size, the comparison is sensitive to the presence of outliers in the data. The majority of outliers stem from Storm 18. Figure 10a presents IMLM, LWM and ECMWF hindcast estimates of spreading over the duration of the storm. In this storm, both  $\sigma_{opt}$  and  $\sigma_E$  follow the same general trend, whereas in the latter half of the storm  $\sigma_D$  presents larger values (about  $20^\circ$  greater). The correlation between  $\sigma_{opt}$  and  $\sigma_E$  may suggest that the IMLM is in error in this case. The apparent error in  $\sigma_D$  may be a result of crossing-wave

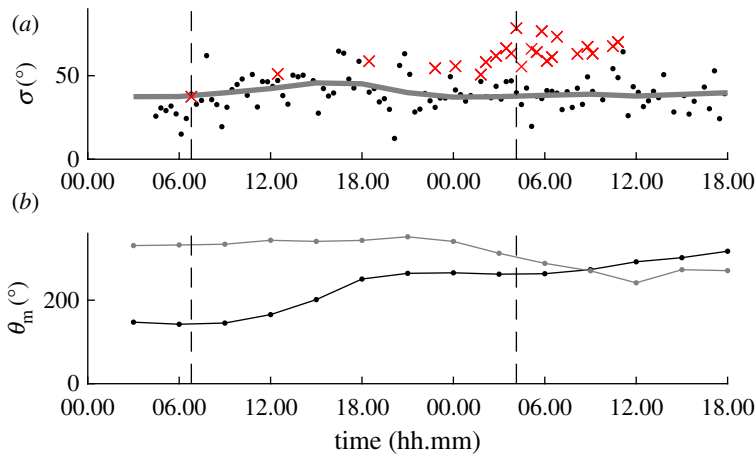


**Figure 9.** Comparison of IMLM and LWM estimates: scatter plot of  $\sigma_{\text{opt}}$  against  $\sigma_D$ , (a) raw  $\sigma_{\text{opt}}$ , and (c) smoothed  $\sigma_{\text{opt}}$ ; and Bland–Altman plot of mean spreading against difference in spreading estimates, (b) raw  $\sigma_{\text{opt}}$ , and (d) smoothed  $\sigma_{\text{opt}}$ . Superimposed mean difference (grey solid line), two standard deviation limits of agreement (grey dashed line), all 386 analysed data (black dots), and the results from storm 18 (crosses). (Online version in colour.)

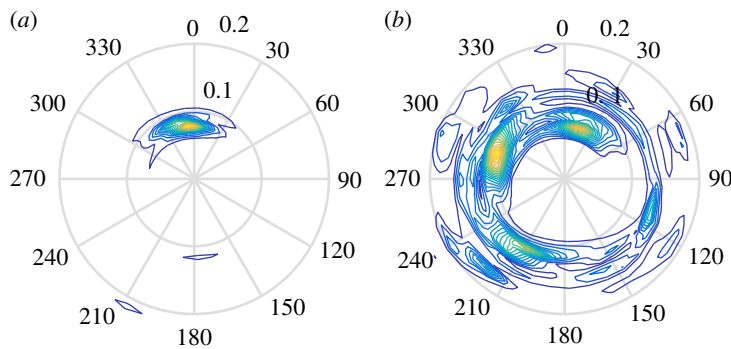
conditions. Figure 10*b* shows the mean directions of wind and swell waves predicted by the ECMWF hindcast for Storm 18. The hindcast results suggest that the predominant wind and swell waves were propagating in quite different directions during the storm. Therefore, the complex crossing conditions are being detected by the frequency-dependent IMLM causing an increase in the value  $\sigma_D$ . The simplicity of the assumed frequency-dependent spreading distribution used by LWM only allows for the detection of the average spreading about the mean direction. This is highlighted by figure 11, which shows frequency-dependent directional spectra produced by the IMLM at the start and one day into the storm, indicated by dashed vertical lines in figure 10. Figure 11*a* shows the spectrum calculated from measurements made at 06.47 on 1 February 2000. At this point in the storm, the spectrum has an obvious predominant direction with no signs of major crossing components, and the spreading estimates made by all three sources agree well, as illustrated in figure 10*a*. Figure 11*b* depicts the spectrum calculated at 04.08 on 2 February 2000. The corresponding spreading estimates shown in figure 10*a* differ significantly ( $\sigma_D = 78.5^\circ$ ,  $\sigma_E = 37.8^\circ$  and  $\sigma_{\text{opt}} = 39.8^\circ$ ), and the directional spectrum exhibits several crossing wind and swell components.

Figure 9 shows that the results from IMLM and LWM have mean bias of  $5.0^\circ$ , and are in close agreement which reduces slightly as spreading increases. The scatter of results indicates there is uncertainty associated with both methods. Figure 12 presents the histograms and kernel density estimates of error between IMLM estimates and both raw, and smoothed LWM results. The histogram in figure 12*a*, for the raw LWM results, displays symmetry with a normal distribution,





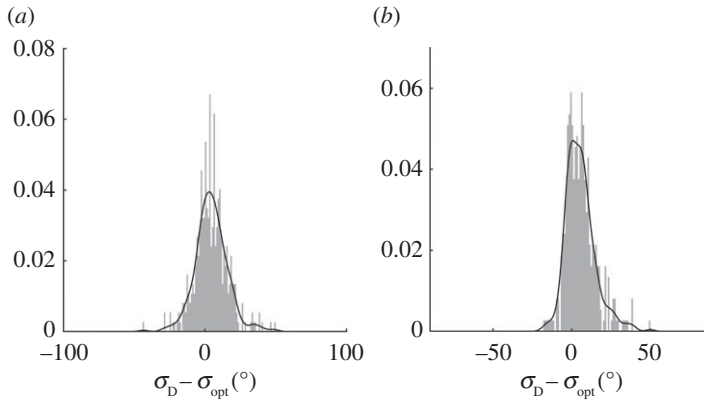
**Figure 10.** Wave spreading results for Storm 18 which exhibit significant disagreement between  $\sigma_{\text{opt}}$  and  $\sigma_{\text{D}}$ : (a)  $\sigma_{\text{opt}}$  (black dots),  $\sigma_{\text{E}}$  (grey line) and  $\sigma_{\text{D}}$  ( $\times$ ); and (b) mean directions of wind waves (black) and total swell (grey). Dashed lines in both plots indicate the measurement times of figures 11*a*,*b*. (Online version in colour.)



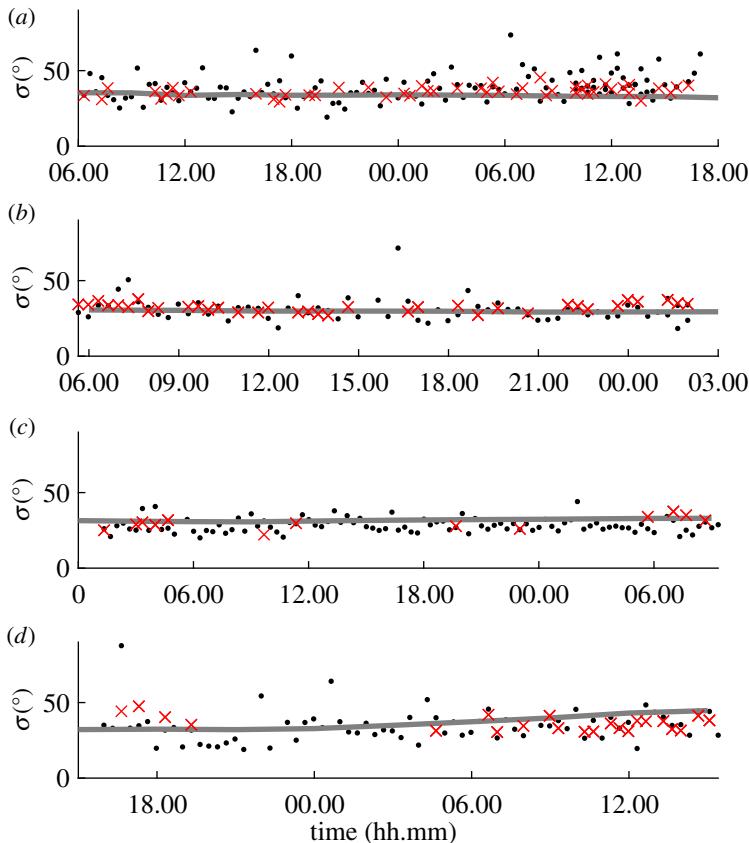
**Figure 11.** Frequency-dependent direction spectrum measured during Storm 18, calculated using IMLM: (a) 06.47 1 February 2000; (b) 04.08 2 February 2000. (Online version in colour.)

suggesting the errors primarily arise from random noise. The histogram in figure 12*b* is slightly asymmetric as a result of smoothing  $\sigma_{\text{opt}}$ , and exhibits better agreement as a narrowing of the distribution.

In the dataset, there are four storms for which IMLM estimation is possible over their duration and a better understanding can thus be gained of the relative uncertainty of the LWM and IMLM methods. Figure 13 displays the four storms detailed in table 4. The first section of Storm 3 was included, and the remainder rejected, in order to preserve data quality and the relative stationarity of spreading. For all four storms,  $\sigma_{\text{opt}}$  (black) and  $\sigma_{\text{D}}$  ( $\times$ ) follow the same trend as the hindcast predictions  $\sigma_{\text{E}}$ , with relatively little temporal variation. These examples further reinforce that LWM has correctly estimated the local spreading given that all three estimates are in good agreement. In figure 13*d*, there is a slight time lag between the values of  $\sigma_{\text{E}}$  and the other two estimates which are better correlated. This storm also shows slightly more variation in spreading, whereas the other storms remain virtually stationary. Table 5 lists values of the mean and standard deviation of the spreading estimates obtained using IMLM and LWM for the four storms. Where the value of spreading remains constant, the standard deviation of each set of estimates provides a satisfactory analogue to uncertainty associated with the particular method employed, assuming that a near-constant value of spreading is a true representation of the actual conditions. Dividing

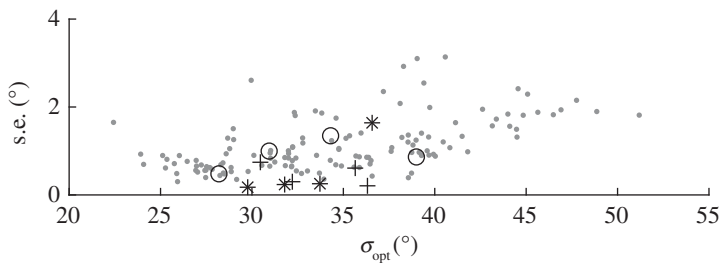


**Figure 12.** Histograms and kernel density of difference in wave spreading estimated by LWM and IMLM methods: (a) raw  $\sigma_{\text{opt}}$ ; and (b) smoothed  $\sigma_{\text{opt}}$ .



**Figure 13.** LWM, IMLM, and ECMWF hindcast predictions of wave spreading for (a) Storm 3, (b) Storm 57, (c) Storm 369 and (d) Storm 448 as detailed in table 4:  $\sigma_{\text{opt}}$  (black dots),  $\sigma_D$  ( $\times$ ) and  $\sigma_E$  (grey line). (Online version in colour.)

the standard deviation by the square root of the number of samples gives the standard error (s.e.), which is a measure of variability accounting for the number of samples [32]. The standard error of  $\sigma_{\text{opt}}$  is roughly two to three times that of  $\sigma_D$  for all but Storm 369 where  $\sigma_{\text{opt}}$  is just over half the value of  $\sigma_D$ . This storm happens to have the lowest spreading value predicted by both methods ( $\sigma_{\text{opt}} = 27.9^\circ$ ,  $\sigma_D = 30.5^\circ$  and  $\sigma_E = 31.8^\circ$ ).



**Figure 14.** Uncertainty measured by standard error versus mean predicted value of spreading for each storm: mean  $\sigma_{\text{opt}}$  (grey dots); storms from table 4  $\sigma_{\text{opt}}$  ( $\circ$ ),  $\sigma_{\text{D}}$  (+) and  $\sigma_{\text{E}}$  ( $\times$ ).

**Table 4.** Details of selected storms with IMLM directional data, showing the number of high-quality time series  $N_{\text{QC}}$  and directional data  $N_{\text{Dir}}$ .

storm	start	end	$N_{\text{QC}}$	$N_{\text{Dir}}$
3*	23 Mar 1997 06.19.14	29 Mar 1997 09.26.32	851	110
57	1 Apr 2000 05.38.48	2 Apr 2000 01.58.58	161	36
369	22 Nov 2002 03.40.26	23 Nov 2002 03.20.38	202	15
448	26 Sep 1996 15.58.30	27 Sep 1996 06.38.30	148	22

**Table 5.** Fit parameters of IMLM storms, standard deviation (s.d.), standard error (s.e.), and mean  $\sigma_{\text{m}}$  for all three estimates of spreading.

storm	s.d.			s.e.			$\sigma_{\text{m}}$		
	$\sigma_{\text{opt}}$	$\sigma_{\text{D}}$	$\sigma_{\text{E}}$	$\sigma_{\text{opt}}$	$\sigma_{\text{D}}$	$\sigma_{\text{E}}$	$\sigma_{\text{opt}}$	$\sigma_{\text{D}}$	$\sigma_{\text{E}}$
3	8.87	3.29	0.919	0.820	0.208	0.255	38.1	36.3	33.7
57	7.90	2.94	0.496	0.987	0.298	0.175	30.6	32.2	29.8
369	4.39	3.93	0.819	0.450	0.744	0.237	27.9	30.5	31.8
448	9.93	4.61	4.92	1.20	0.610	1.64	33.5	35.7	36.6

Making the assumption that the smoothed value of  $\sigma_{\text{opt}}$  for each storm is the actual local spreading, the data can be de-trended and a similar analysis performed on the entire dataset. Figure 14 shows how the standard error of  $\sigma_{\text{opt}}$  varies with the mean value of  $\sigma_{\text{opt}}$  over the duration of each storm plotted as grey dots. The values for the four storms plotted in figure 13 and listed in table 4 are overlaid. There is a clear positive correlation between the predicted spreading and the standard error. This is an intuitive result because the simplified assumption of frequency-independent wrapped-normal spreading becomes less valid as spreading increases.

As discussed in §5a, there are a number of estimates  $\sigma_{\text{opt}} < 20^\circ$  which are not within the range of  $\sigma_{\text{E}}$ . Figure 9a also shows values of  $\sigma_{\text{opt}} < 20^\circ$ ; however, after smoothing the number reduces (figure 9c). Figure 14 shows that no storms with multiple concurrent measurements have a mean value of  $\sigma_{\text{opt}} < 20^\circ$ , therefore these values are most likely the result of noise.

6. Conclusion

Adcock & Taylor [14] devised a means of determining directional spreading, the LWM, which they verified for several deterministic cases and later applied to the 1 January 1995 Draupner wave

[16]. Herein, we extend the approach of the LWM to a large dataset of *in situ* ocean observations of free surface elevation obtained from the North Alwyn platform in the northern North Sea. The approach taken to implement the LWM, greatly reduced its computational cost. The LWM analysis was based on an assumption that directional spreading was independent of frequency, and excluded records containing freak waves. A method for detecting the influence of erroneous measurements is also presented, which allows *a posteriori* identification of spurious values of estimated spreading. The results of the LWM are compared with values predicted by the ECMWF hindcast model using the North Alwyn data. Despite the indirect nature of this comparison, good temporal correlation is found between the two estimates of spreading for selected storms. This provides confidence in the ability of the LWM to track dynamically varying directional spreading conditions. By comparing results using a small sample of the overall dataset, it is found that close agreement is achieved between the spreading estimates made by LWM and the IMLM, with the former giving smaller values indicating a slight bias. The LWM exhibits slightly greater uncertainty than the IMLM; however, given that LWM requires only a single measurement, the uncertainty can be reduced by averaging. To gain better understanding of the small bias that exists between the two measurement approaches and the performance of LWM at values of  $\sigma < 20^\circ$ , it is recommended to extend the analysis in future to a dataset containing a much higher proportion of accurate, concurrent measurements. This would enable a quantitative assessment of the relative uncertainty of both methods. This study has demonstrated that the LWM can be effectively extended to noisy real sea observations, such as prevail at North Alwyn. Available additional information has been used to provide confidence in the accuracy of the results. As with the conventional IMLM method, the LWM is susceptible to the effects of noise which manifests itself in the form of random uncertainty. However, where sufficient results exist, averaging effectively removes uncertainty without requirement of further post-processing. In short, the study has provided further confirmation that the LWM is a viable alternative to conventional means of directional spreading estimation. In practice, LWM offers a very promising opportunity to gain vital directional information on ocean waves from single-point measurements.

**Data accessibility.** The datasets used herein are owned and distributed by Total S.A, and the ECMWF.

**Authors' contributions.** M.L.M. implemented the theory and performed the analysis of the North Alwyn dataset, with the supervision and guidance of V.V. and A.G.L.B.

**Competing interests.** We declare we have no competing interests.

**Funding.** The first-named author is the recipient of an EPSRC studentship under the DTP.

**Acknowledgements.** The authors would like to thank Prof. Tom Adcock and Prof. Paul Taylor of the University of Oxford who provided useful advice and access to their potential flow results. The authors are grateful to Total S.A. for the wave data from the North Alwyn platform and ECMWF for the hindcast data.

## References

1. Benoit M, Frigaard P, Schäffer HA. 1997 Analysing multidirectional wave spectra: a tentative classification of available methods. In *Proc. IAHR Seminar on Multidirectional Waves and their Interaction with Structures, 27th IAHR Congress, San Francisco, CA*, pp. 131–154.
2. Forristall GZ, Ewans KC. 1998 Worldwide measurements of directional wave spreading. *J. Atmos. Oceanic Tech.* **15**, 440–469. (doi:10.1175/1520-0426(1998)015<0440:WMODWS>2.0.CO;2)
3. Mitsuyasu H, Tasai F, Suhara T, Mizuno S, Ohkusu M, Honda T, Rikiishi K. 1975 Observations of the directional spectrum of ocean waves using a coverleaf buoy. *J. Phys. Oceanogr.* **5**, 750–760. (doi:10.1175/1520-0485(1975)005<0750:OOTDSO>2.0.CO;2)
4. Barstow SF, Kallstad T. 1991 Field trials of the directional waverider. In *Proc. of the First International Offshore and Polar Engineering Conference, Edinburgh, UK, 11–16 August*, vol. III, pp. 55–63. Mountain View, CA: ISOPE.
5. Walsh EJ, Hancock III DW, Hines DE, Swift RN, Scott JF. 1985 Directional wave spectra measured with the surface contour radar. *J. Phys. Oceanogr.* **15**, 566–592. (doi:10.1175/1520-0485(1985)015<0566:DWSMWT>2.0.CO;2)
6. Stokes GG. 1847 On the theory of oscillatory waves. *Trans. Camb. Philos. Soc.* **8**, 441–473.

7. Longuet-Higgins MS. 1963 The effect of non-linearities on statistical distributions in the theory of sea waves. *J. Fluid Mech.* **17**, 459–480. (doi:10.1017/S0022112063001452)
8. Johannessen TB, Swan C. 2001 A laboratory study of the focusing of transient and directionally spread surface water waves. *Proc. R. Soc. Lond. A* **457**, 971–1006. (doi:10.1098/rspa.2000.0702)
9. Onorato M *et al.* 2009 Statistical properties of mechanically generated surface gravity waves: a laboratory experiment in a three-dimensional wave basin. *J. Fluid Mech.* **627**, 235–257. (doi:10.1017/S002211200900603X)
10. Forristall GZ. 1999 Wave crest distributions: observations and second-order theory. *J. Phys. Oceanogr.* **30**, 1931–1943. (doi:10.1175/1520-0485(2000)030<1931:WCDOAS>2.0.CO;2)
11. Toffoli A, Monbaliu J, Onorato M, Osborne AR, Babanin AV, Bitner-Gregersen E. 2007 Second-order. *J. Phys. Oceanogr.* **37**, 2726–2739. (doi:10.1175/2007JPO3634.1)
12. Walker D, Taylor PH, Eatock Taylor R. 2004 The shape of large surface waves on the open sea and the Draupner New Year wave. *Appl. Ocean Res.* **26**, 73–83. (doi:10.1016/j.apor.2005.02.001)
13. Dalzell JF. 1999 A note on finite depth second-order wave-wave interactions. *Appl. Ocean Res.* **21**, 105–111. (doi:10.1016/S0141-1187(99)00008-5)
14. Adcock TAA, Taylor PH. 2009 Estimating ocean wave directional spreading from an Eulerian surface elevation time history. *Proc. R. Soc. A* **465**, 3361–3381. (doi:10.1098/rspa.2009.0031)
15. Cornett A, Morin V, Durand N. 2002 Model tests to select deck elevations for the PA-B and Lun-A platforms. Technical Report HYD-CTR-008. Tech. rep., Canadian Hydraulics Centre.
16. Adcock TAA, Taylor PH, Yan S, Ma QW, Janssen PAEM. 2011 Did the Draupner wave occur in a crossing sea? *Proc. R. Soc. A* **467**, 3004–3021. (doi:10.1098/rspa.2011.0049)
17. Haver S. 2004 A possible freak wave event measured at the Draupner Jacket January 1 1995. In *Rogue Waves 2004, Brest, France, 20–22 October*, pp. 1–8. Issy-les-Moulineaux, France: Ifremer.
18. Rosenthal W, Lehner S. 2008 Rogue waves: results of the MaxWave project. *J. Offshore Mech. Arct.* **130**, 21006. (doi:10.1115/1.2918126)
19. Sawaragi T. 1995 Chapter 2 Wave interactions with structures and hydrodynamic forces. *Dev. Geotech. Eng.* **78**, 67–149. (doi:10.1016/S0165-1250(06)80037-4)
20. Christou M, Ewans K. 2011 Examining a comprehensive dataset containing thousands of freak wave events: Part 1—description of the data and quality control procedure. In *ASME 2011 30th International Conference on Ocean, Offshore and Arctic Engineering*, pp. 815–826. New York, NY: American Society of Mechanical Engineers.
21. Wolfram J, Feld G, Allen J. 1994 A new approach to estimating environmental loading using joint probabilities. In *7th International Conference on Behavior of Offshore Structures*, pp. 701–704. Boston, UK: Pergamon.
22. Christou M, Ewans K. 2014 Field measurements of rogue water waves. *J. Phys. Oceanogr.* **44**, 2317–2335. (doi:10.1175/JPO-D-13-0199.1)
23. ECMWF. 2016 IFS documentatation Cy40r1 Operational implementation, Part VII: ECMWF Wave Model. Technical Report. Reading, UK, 1–79.
24. Gibbs RH, Taylor PH. 2005 Formation of walls of water in ‘fully’ nonlinear simulations. *Appl. Ocean Res.* **27**, 142–157. (doi:10.1016/j.apor.2005.11.009)
25. Bateman W, Swan C, Taylor PH. 2001 On the efficient numerical simulation of directionally spread surface water waves. *J. Comput. Phys.* **174**, 277–305. (doi:10.1006/jcph.2001.6906)
26. Haver S. 2001 Evidences of the existence of freak waves. In *Rogue Waves, Brest, France, 29–30 November 2000*, pp. 129–140. Issy-les-Moulineaux, France: Ifremer.
27. Tromans PS, Anatrak AR, Hagemeijer P. 1991 New model for the kinematics of large ocean waves application as a design wave. *Proc. First Int. Offshore Polar Eng. Conf.* **8**, 64–71.
28. Santo H, Taylor PH, Eatock Taylor R, Choo YS. 2013 Average properties of the largest waves in Hurricane Camille. *J. Offshore Mech. Arct.* **135**, 0116 021–0116 027. (doi:10.1115/1.4006930)
29. Adcock TAA, Draper S. 2015 The second order contribution to wave crest amplitude—random simulations and New Wave. In *The Twenty-fifth International Offshore and Polar Engineering Conference, Kona, HI, 21–26 June*. Mountain View, CA: ISOPE.
30. Emery WJ, Thomson RE. 2001 Chapter 3 - Statistical methods and error handling. *Data Anal. Methods Phys. Oceanogr.* 193–304.
31. Bland JM, Altman DG. 1999 Measuring agreement in method comparison studies. *Stat. Methods Med. Res.* **8**, 135–160. (doi:10.1191/096228099673819272)
32. Thiébaux HJ. 2013 *Statistical data analysis for ocean and atmospheric sciences*. Amsterdam, The Netherlands: Elsevier.